图书情报工作
第 64 卷 第 3 期 2020 年 2 月
ChinaXiv合作期刊

欧盟大数据伦理治理实践及对我国的启示*

■ 陈一
武汉大学信息管理学院 武汉 430072

摘 要：[目的/意义]总结欧盟大数据伦理治理的经验，为我国大数据伦理治理提供参考。[方法/过程]系统梳理欧盟有关大数据伦理治理的相关政策，分析欧盟大数据伦理治理的实践，重点剖析欧盟社会经济委员会(EESC)的治理措施，在此基础上总结欧盟的治理经验，并根据我国的实际情况，归纳出对我国大数据伦理治理的启示。[结果/结论]EESC 从人类生命周期出发，将大数据伦理归纳为包括“意识”“控制力”等在内的 10 个主要问题，并从个人、企业、研究机构多层面提出 5 项治理举措，同时强调从多视角审视大数据伦理问题。而目前我国数据伦理法律规范的滞后，以及公民伦理意识的缺失给大数据利用带来了巨大的伦理风险，我国应借鉴欧盟治理经验，明确大数据环境下的个人权利，在数据治理过程中关注多利益主体的伦理需求，构建数字教育体系，探索建立研究者持续性审核责任机制。

关键词：大数据 伦理 治理
分类号：B829
DOI：10.13266/j.issn.0252-3116.2020.03.014

chinaXiv:202304.00343v1

1 引言

大数据正在对社会各领域、各行业产生深刻影响，运用大数据推动经济发展、完善社会治理、提升政府服务和监管能力已经成为趋势。正是因为数据成为重要的战略性资产，其带来的巨大社会和经济效益不可避免的导致非法采集、传播、滥用数据等伦理问题，如个人数据非法获取和保存、数据滥用、数据主体对数据的控制权被削弱、数据垄断、数据不公平应用、数据偏向引导等。大数据与智能技术的应用也伴随着复杂风险，如数据计算非法预测、获取他人身份等。*Science* 2018 年 8 月发表综述文章 *How AI can be a force for good*[1]。文中表示，人工智能正在重塑日常实践、个人与专业之间的互动以及环境，将这种力量用作善的力量是至关重要的。伦理规范在这一过程中发挥着关键作用，它能够确保充分发挥人工智能潜力的同时降低风险。AI 受到数据的推动，因此面临与数据治理相关的道德挑战，包括同意、所有权、隐私等，AI 可能会加剧这些与数据相关的挑战。

伦理是处理各种发展关系时应坚持的道德准则，数据伦理问题是关注在收集、分析数据过程中，以及在生物医学研究和社会科学研究中使用、描述、传播及开放数据等一系列活动中所产生的伦理问题[2]。在参考相关学者的理论研究后[3-4]，本研究认为大数据伦理治理则是不同的人群或组织机构在大数据时代为了应对在大数据利用过程中带来的伦理失范而运用不同的政策、技术工具对大数据采集、利用、存储等过程进行管理、规范并挖掘其价值的行为，治理过程也不再只是出于政府的行政权力，而是多方协调所达成的一致。规范大数据利用过程中的伦理问题对正处在社会转型时期的我国利用大数据推进社会治理具有重要的理论与现实意义。

近年来，欧盟有关“数字化”的观点呈指数级增长，对伦理层面有很强的前瞻意识，广泛开展大数据伦理调查[5]。随着大数据、人工智能的深度发展，欧盟对其中伦理的规范也更为严格。2018 年 4 月欧盟发布政策文件《人工智能》[6]（*Artificial Intelligence*），其中将“确保欧盟具有与人工智能发展及应用相适应的伦理和法律框架”作为人工智能战略发展的三大支柱之一；2018 年《欧洲通用数据保护条例》（*General Data Protection Regulation*，简称 GDPR）[7]的颁布规定了大数据时代的伦理红线和考量原则；“人工智能欧洲造”（AI

* 本文系中国博士后科学基金资助项目“大数据应用中的伦理风险演化及其治理框架构建”（项目编号：2017M622526）研究成果之一。
作者简介：陈一（ORCID：0000-0003-1828-1165），博士后，E-mail：chenyi@whu.edu.cn。
收稿日期：2019-03-27 修回日期：2019-09-24 本文起止页码：130-138 本文责任编辑：徐健

made in Europe)[8]计划中两大关键原则首位便是"设计伦理"(ethics by design);2019 年《值得信赖的人工智能伦理准则》(*Ethics Guidelines for Trustworthy Artificial Intelligence*)[9]的发布进一步显示了欧盟对保护个人面对技术、数据的选择权、知情权的重视。

可见,欧盟在大数据环境下,对于数据及技术使用伦理的监管一直走在世界前列。本文系统调研欧盟大数据伦理治理的相关政策,总结欧盟大数据伦理的治理实践,并通过追溯检索尽可能保证调查文本全面性。对欧盟相关经验的总结更有助于深刻理解我国大数据开发利用中的伦理问题,为我国大数据伦理治理提供有益的参考。

2 国内外大数据伦理研究进展

2.1 大数据伦理问题的具体表现

2012 年 9 月,美国学者 K. Davis 和 D. Patterson[10]出版了国际上首部有关大数据伦理的学术专著 Ethics of Big Data,作者认为,大数据环境下,企业应该确立自身的道德规范,明确数据对于自身的价值,重视数据中所涉及的身份(identity)、隐私(privacy)、归属(ownership)以及名誉(reputation),在技术创新与风险之间寻求平衡。英国皇家学会发布的报告中认为,在一些调查和公共对话中受关注的伦理问题可归纳为:隐私、自由、自治和自决,其中一些问题也特别针对机器学习等新技术的使用[11]。美国华盛顿大学法学教授 N. Richards 等认为大数据伦理应坚持 4 项原则:①将"隐私"作为数据规则;②共享私人信息也应保持"机密";③数据利用应有"透明"的机制;④要意识到数据可以破坏"身份"[12]。2016 年牛津互联网学院信息哲学与伦理学教授、数字伦理实验室主任 F. Luciano 和 T. Mariarosaria 博士组织专题探索大数据伦理究竟是什么(what is data ethics?),他们认为其中主要包括 3 个领域:①数据的伦理(如何生成、记录和共享数据);②算法的伦理(人工智能、机器学习和机器人如何解释数据);③实践的伦理(设计负责任的创新和专业守则来指导这一新兴科学)[13]。

我国学者薛孚等[14]认为大数据推动社会向前发展的同时引发了不同于以往的隐私伦理挑战,这种新挑战表现在数据挖掘、数据预测和更全面的监控等方面。邱仁宗[15]教授以数字身份为核心,总结了大数据技术在创新、研发及应用中的存在的伦理问题,其中"数字身份""个人隐私""数据可及""数据安全""数字鸿沟"等伦理问题尤其值得关注。安宝洋[4]认为,大数据在给人类社会带来积极变化的同时,也引发了信息异化、数据权利、信息隐私和数字鸿沟等网络信息伦理问题。

2.2 大数据伦理问题成因

尽管大数据给全球带来如此巨大的影响,但实际上,社会、政府及法律体系都还未真正适应大数据时代的到来[16]。目前,国内外学界一致认为,数据本身是一种客观存在,大数据本身没有好与坏之分,产生伦理问题的根源是人类对数据的不合理使用。大数据巨大的商业价值使得将个人数据转向可分析的数据,以获得更多的产品和利润[17],而商业的介入则催生了道德风险,现代网络信息技术的完善成熟又为这种利益追逐提供了更为自由和畅通的选择空间,大大增加了预警和规避伦理风险的实际操作难度。

也有不少学者从大数据应用的负效应和法律失范和伦理道德滞后[18]等方面具体分析了伦理问题产生的原因。"数据需要管理,但问题在于人们正在逐步摆脱数据收集的管理"[19]。王小杨[20]认为,大数据陷入伦理困境有 4 个方面的主要原因:大数据世界观的缺位、数据权利边界的模糊、数据市场应用的失衡以及大数据技术开发的短视,其根源在于人的主体性弱化与大数据语境中伦理边界失守。陆伟华[21]认为伦理问题产生来源于五大冲突:数据采集、应用与个人隐私的冲突;大数据管理与信息安全的冲突;大数据时代知识产权观念的变革与现有知识产权保护的冲突;大数据信息滥用与信息危机的冲突;大数据信息污染与信息生态环境的冲突。赵素锦[22]从利益主体的角度思考,认为每一个个体的行为活动都是整个高科技活动链条上的一个环节,每一个人几乎都是伦理风险的直接或间接制造者,这使得伦理风险的制造主体处于"漂流"中的不确定状态,而这种不确定使得责任主体难以判定,扩大了伦理问题的危害。

2.3 大数据伦理问题治理策略

信息技术的发展使伦理问题陷入困境,且日益复杂,学者们也迫切希望探索出适用的伦理准则与规范。L. Judd[23]早在 1995 年就提出了信息时代应坚守的伦理与道德准则:①适当承担责任;②预料消极影响;③以约翰· 罗尔斯(J. Rawls)的正义原则为指导追求公平;R. Severson[24]倡导利用尊重知识产权、尊重隐私、公平参与及无害这四原则解决信息伦理问题;一些相关的学术组织也制定了相关的伦理道德规范,如电气和电子工程师协会[25]、美国计算机协会[26]、美国统计协会[27]、英国计算机学会[28]等都制定了相关的伦理

chinaXiv:202304.00343v1

ChinaXiv合作期刊

行为准则。伦理规则在促进行业自律和政策制定、规范信息人员行为和帮助人们进行伦理决策方面,具有明显的积极意义[29],这些也都给大数据伦理治理提供了有益的借鉴。

舍恩伯格在其著作《删除——大数据时代的取舍之道》[30]中认为,在这个遗忘成为例外,而记忆成为常态的环境中,要把有意义的留下来,无意义的去掉。A. Zwitter[31]提出了在法律和政治领域的发展空间。D. Nunan 等[32]建议要贯彻被遗忘权、数据到期权、社交图片所有权这 3 项伦理实践,来保障个人对自身的隐私控制。我国学者于英香[33]认为,需要构建一个全面的大数据治理体系,该体系包括战略、规划等宏观领域,制度、组织、流程等中观领域,软硬件平台、标准、法规、技术等微观领域。宋吉鑫[34]认为应该采用大数据技术制约、法制规范与自律教育相结合的方式,提高网民网络素养,规范网民、企业的网络行为。陈万求[35]也认为应该建立起一套法律、伦理、监督相结合的机制调控大数据技术的发展。

2.4 研究进展述评

通过对文献的梳理发现,目前研究多是对大数据环境下伦理问题的笼统归纳,而大数据开发利用是一个动态演化的过程,数据主体也具有多样性,不同主体对大数据伦理问题的认知也存在差异,目前的研究也没有将利益相关者考虑在内。当下有关大数据伦理的治理策略,还大多停留在呼吁阶段,也多为末端治理,实践中对大数据伦理问题的规范效果并不理想。欧盟在技术应用与伦理层面一直都具有很强的前瞻意识,本文系统梳理了欧盟大数据伦理治理相关政策及实践的探索,希望能为我国大数据伦理治理提供切实有用的参考,建立引导型的行为规范和预防机制,实现对大数据伦理问题的前瞻性规避。

3 欧盟大数据伦理治理进程

3.1 欧盟大数据伦理治理相关政策演进

数据已经成为经济发展和社会进步的一项关键资产,欧盟在抓住大数据提供的机遇同时,并没有忽视其伴随的挑战。近年来,欧盟发布多项战略、计划,颁布相关政策,要求在利用大数据时保障个人权利,关注数据权属、数据安全、数据鸿沟等问题。

2012 年 9 月,欧盟委员会发布了《释放欧洲云计算的潜力》(*Unleashing the Potential of Cloud Computing in Europe*)[36],计划 2014 - 2020 年之间推动成员国利用云计算的潜力,同时表示委员会会尽快解决一般网络安全挑战,采取适当的技术和组织措施来管理安全风险,并向主管当局报告重大事件的风险。

2013 年 2 月,欧盟委员会发布了欧洲网络安全战略《一个开放、安全、可靠的网络空间》(*An open, safe and secure Cyberspace*)[37],其中强调欧洲当局应该关注网络空间中需要保护的基本权利——民主与法治,欧盟必须促进“网上自由”,并确保“尊重在线基本权利”,这可以通过授权访问权限、建立民主和有效的多主体治理体系以及责任分担来实现。

2013 年 3 月,欧盟委员会发起了数字技能与就业联盟(The Digital Skills and Jobs Coalition)[38],该计划旨在提升所有人的数字技能,使所有公民都能参与到数字社会;提升劳动力的数字职业技能,加强职业建议和就业指导;提升 ICT 专业人员高水平的职业技能;转变数字技能的教育体系,建立终身学习体系。该计划能有效提高欧盟公民的数字素养,有效弥合数据鸿沟。

从更务实的层面看,欧盟委员会 2013 年 12 月批准实施地平线 2020(Horizon 2020)科研计划,“大数据”是这一计划的核心,对大数据相关研究提供直接经费支持,也鼓励从创新的角度探索大数据伦理问题。

2015 年 11 月,欧洲数据保护主管撰写了 7/2015 号意见[39],就大数据本身的风险和挑战展开了讨论,呼吁透明度、用户控制和数据保护,还强调对大数据的开发利用必须依赖 4 要素:组织必须利用更透明的方式来处理个人数据;用户能更好地掌控个人数据;在产品和服务中设计更友好的数据保护;担负起对数据开发利用的责任。

2018 年 5 月 25 日,《欧洲通用数据保护条例》[7]正式实施。这个被称为欧盟最严格的数据保护法案强调个人数据权利,不再将“个人权利”作为数据保护原则的一条,而是运用一整章单独规范,确立了包括知悉权、访问权、整改权、擦除权等在内的 8 项个人权利,使人们能够更有效地控制自己的数据。

2019 年 4 月 8 日,欧洲委员会人工智能高级专家组(High Level Group on Artificial Intelligence, AI HLEG)发布了《值得信赖的人工智能伦理准则》[8],旨在最大限度发挥人工智能作用,同时最大程度地降低风险。其中规定,AI 系统开发利用中应坚持下列原则:有益(do good)、无害(do no harm)、保持人的能动性(preserve human agency)、公平(be fair)以及透明操作(operate transparently)。

3.2 欧盟大数据伦理治理实践

针对大数据伦理问题,欧盟也设计了具体的治理

chinaXiv:202304.00323v1

举措。2016 年 6 月,欧盟社会经济委员会(European Social and Economic Committee,EESC)启动了大数据伦理调查[5],调查于 7 月正式开始,12 月结束。调查明确了大数据伦理具体包含哪些问题,并在此基础上制定了相应的制衡措施。而在大数据环境下,各类主体都可能既是数据的生产者,也是采集和利用者,数据治理应实现多元化协同配合,EESC 也意识到必须从多利益主体视角审视这一问题。

3.2.1 10 个问题

明确大数据伦理问题的具体表现是实现治理的首要步骤。EESC 从人类生命周期视角梳理了人们从出生之前到死亡各个阶段会遇到的伦理困境,并将其归纳为 10 个主要问题。

(1)意识(awareness),人们在创建数字身份时无意识地提供了很多个人数据,而当人们利用数字身份访问第三方资源时也可能同时透露了很多个人行为数据而不自知,丧失了必要的知情权。

(2)控制力(control),当用户决定部分或者全部删除他们之前提供给服务商的数据时,即便服务商按照用户需求删除数据,但已经出售给其他公司或已经大批量处理的数据并不会受到影响,人们对自身数据并没有实际控制权[40]。

(3)信任(trust),这是用户愿意提供个人数据的基础,它与隐私和意识相互依存,然而目前人们并未与计算机环境建立起信任关系,当下也主要是通过严格的技术手段来解决信任问题。

(4)所有权(ownership),主要指经过处理后的原始数据面临着复杂的所有权关系,原始的用户数据经过加工后应该如何处理,这些数据是否还是用户数据,或是应该属于执行分析的公司或收集数据的公司? 用户、数据分析公司及数据收集公司之间权利归属存在争议,数据所有权目前没有更为实际的法律和政策解决路径。

(5)监视与安全(surveillance and security),由于数据源的增加和分析技术的进步,能够非常便捷地利用数据生成有价值的信息,利用某种方式追踪监视某人或推测某人的立场也变得很容易。

(6)数字身份(digital identity),人们可以便捷的利用数据身份获取网络服务,但数字身份的广泛使用也使得个人公开可用的信息能够被广泛检索到,人们常常会基于数字身份提供的数据对某人进行评价,而这很可能会造成歧视,即"我们不再根据行为来判断,而是根据数据来判断行为"[41],人之间的交互通常也被放在了分析数字身份之后。

(7)定制的现实(tailored reality),人们通过搜索引擎进行检索的关键词,或在线购物等搜索偏好数据都有可能被存储,对该数据的分析和处理可用于后期根据用户偏好为其提供相关信息,用户正在经历更个性化却更狭窄的在线体验[42],即"过滤泡沫"。

(8)去匿名化(de-anonymization),传统的匿名化技术主要通过删除(替代)唯一可识别信息使每个条目数据不可识别,但这在当今已经无效,各种数据关联分析后能产生强大的洞察力,某种程度上依然能够识别某些个体身份。

(9)数字鸿沟(digital divide),一些用户因为缺乏相关信息技能,难以通过互联网等新技术获取服务,或由于不熟悉这些流程而难以真正理解这些流程的工作方式,如老年人在求职时就经常面临这样的困境。

(10)隐私权(privacy),隐私是一个包罗万象的话题,它涵盖了上述伦理问题的内涵,其主要指人们拥有个人信息非经许可他们不可使用的权利,这是在当前最重要的一个伦理议题。

3.2.2 5 项举措

在对大数据伦理问题系统总结的基础上,EESC 从多角度提出了 5 项治理举措,以期能够全面涵盖上述 10 个问题,从个人、企业、研究机构等各个层面实现有效治理。

(1)欧盟隐私管理平台。其想法是建立一个泛欧洲的门户网站作为隐私管理中心。欧洲公民自愿注册,在上面有个人数据控制中心,公民通过控制中心能够可视化地看到所有存储、处理、再利用这些数据的公共和私人实体列表,并可以自主操控。正是由于大数据环境下用户对数据采集的"无意识""无控制",侵犯隐私屡屡发生,这一平台直接赋予公民控制其个人数据的权利,并提供便捷的退出服务,用户可以决定是否分享、重用个人数据,也能够有效提升信息安全。这也与 GDPR 中强调自然人应控制其个人数据的原则一致。

(2)发布《数据伦理管理协议》。旨在提高透明度,使人们了解大数据持有者(公共和私人)遵守欧盟法律的程度。其理念是设计一个完善的欧洲认证体系,在数据保护领域进行企业认证。经过数据伦理认证的企业显示了对法律和公民权利的尊重,能更好地获得用户"信任",也进一步提升了公司的声誉和经济效益。

(3)发布《数据管理声明》。这一项是建立在公司

chinaXiv:202304.00343v1

ChinaXiv合作期刊

自愿参与的基础上，其出发点在于当今组织的成功越来越依赖于股东、客户、员工和公众的多方信任，因此，为了增强内外部利益相关者的信心，组织可以提交有关“如何收集、使用或销售来自客户或一般业务活动中的数据的声明”。其中要求阐明数据性质、未来用途、使用策略等，并对加密和匿名化技术的相关情况做明确说明。从公众角度看，《数据管理声明》能提高个人数据利用过程的透明度，防止机构过度收集数据，也能够限制第三方对数据或用户数字身份的不道德使用，加强数据的主体控制。

（4）建立欧洲电子健康数据库。对公民健康数据的规范利用是保障大数据伦理的关键领域。该措施计划建立包含欧洲公民健康数据相关的欧洲数据库，用于科学研究。当欧盟公民接受公共资助医疗时，会被询问是否同意数据被收集并存储于该数据库中，个人同意后，数据会按照标准交换协议被收集和传送。学者和研究机构必须提交申请才可利用数据，申请除了包括研究者及研究项目的具体信息之外，还要求说明数据申请理由、利用方式及预期结果。通过评估的学者或研究机构会得到被匿名化和粗化处理（其程度达到防止数据去匿名化）的数据。同时，公民个人可以利用其法定数字身份信息访问数据库，可对个人医疗数据进行管理和变更。这种规范的医疗数据利用途径能够大大降低非法采集和利用医疗数据的概率，有效促进科学研究，同时提高欧洲公民对自身医疗数据的掌控权，增加对医疗健康服务与研究的信任，也丰富欧洲公民的数字身份信息。

（5）构建大数据时代的数字教育体系。这一举措旨在在欧洲建立更广泛的数字文化，使欧洲公民对大数据有更深入的理解，使其明白大数据如何在生命周期内与欧洲公民互动并影响到每个人。为提升这一意识，要针对不同年龄段人群设计教育计划。小学至高中主要以介绍大数据为主，包括大数据基本概念，什么是数字身份，如何收集数据，怎样避免过度泄露个人隐私等。大学学位教育主要以培养相关专业人才为主，通过系统学习，创建和使用大数据的伦理学方法。在社区中，按照不同年龄层，特别是针对老年人的需求，举办不同的专题，提升居民的数字化技能。数字教育体系的构建能够有效提高欧洲公民的数字素养，缩小数字鸿沟，也能够对大数据环境有更清晰的认知，对网络环境所提供的信息有更理性的认知。另外，也能对当今“享受数字服务可能需要牺牲部分隐私”这一风险有一定预判，以增加人们对新技术的接受度和信任度。从长期看，大数据伦理意识的培养，使人们有意识地在伦理框架下收集、利用数据，也有助于数字社会的健康发展。

chinaXiv:202304.00343v1

3.2.3 多个视角

不同的个人、组织或机构处在数据链条的不同位置上，视角不同，对大数据伦理问题的认知也存在差异。为了获取广泛的意见，EESC 进行了网络问卷调查，并从数据保护机构、大数据分析公司、中小企业、消费者协会、欧盟机构、学术界和数据科学专家中选出了 16 位代表进行了深入访谈，虽然大家都表现出对目前大数据伦理现状的担忧，但大家具体的态度却存在差异，学术界和国家数据保护机构主要关注的是数据合理使用、透明度等有关大数据伦理的一般性议题，他们认为过度个性化的数字体验，有可能形成偏见，而有 2 位数据专家则对数据利用保持乐观态度，并强调了大数据产生新见解的潜力。

受访者对 5 项举措也持有不同的态度，例如数据专家认为发布《数据伦理管理协议》，建立认证体系可能会增加公司的负担，而受访者中唯一一位公司经理却并不认为这是额外的负担。相反的，他认为这是提高用户对他们公司数据使用意识的好工具。针对建立欧洲电子健康数据库，绝大多数受访者虽然都认可这将对医疗和研究都提供很大帮助，但是并不认为这一举措在当下是可行的，主要是出于对敏感数据匿名化和潜在安全性的担忧。受访者对数字教育计划的态度非常一致，认为这一措施在信息社会中发挥重要作用，应尽早实施。

3.3 欧盟大数据伦理治理经验

近年来，欧盟在利用大数据创造价值的同时，也始终关注着伦理问题，不仅仅是呼吁号召，相较于其他对大数据伦理问题的“案例型”总结，欧盟对大数据伦理的梳理则更加全面，并从多维度推行治理举措，形成了一些的经验值得借鉴。

3.3.1 从“人”的角度出发

大数据伦理问题归根到底与“人”相关，人们甚至在出生之前就开始参与到数字世界中，并在整个生命周期阶段以不同方式和不同意识水平继续提供和利用着数据。从欧盟各项政策中对“民主”“透明”“个人权利”等的强调，到 EESC 基于人的生命周期凝练伦理问题，推行具体举措保障个人隐私数据，再到 AI HLEG 所强调的“隐私和数据治理”“多样性、不歧视和公平”，关注不同立场的伦理需求。而且在 2019 年 4 月 8 日，欧盟委员会与议会、理事会、欧盟经济社会委员会

以及区域委员会进行沟通,表示要建立"以人为中心(human centric)"的人工智能信任[43],这些都体现出对"人"的关注,对个人权利的重视。

3.3.2 多视角推行治理举措

大数据环境下,各类主体都可能既是数据的生产者,也是采集和利用者,不同主体对大数据伦理问题的认知也存在差异,对规范大数据伦理也有不同程度的需求。有调查表示 46% 的英国成年人认为他们"被大公司收集的数据所伤害"[44],但另一研究发现,60% 的人"对与公司共享个人信息量感到满意",47% 的人认为"个人数据的交换对于现代社会的顺利运行至关重要"[45]。欧盟的多项治理举措从个人、企业、研究机构等各个层面出发,考虑不同主体立场,并重点关注到了面临严峻伦理问题的健康数据领域。2019 年 6 月在布鲁塞尔举行的第一届 AI 联盟大会上,AI HLEG 也表示要针对市场量身定制伦理政策,促进社会、行业组织、公共部门以及学术界利益相关者的合作[46],不再是政府单一、单向监管,而是多主体协同,以有效推进治理进程。

3.3.3 重视数字教育

欧盟非常关注居民数字素养的提升。欧盟委员会发起的数字技能与就业联盟旨在提升所有人的数字技能,EESC 倡导构建大数据时代的数字教育体系,社区与之联动,通过各类型培训使各年龄层居民对数字社会有更全面的认知,培养数据伦理意识。提高对数字社会的认知水平可以有效提高公民的参与水平,从而在尊重公民尊严的伦理背景下为政府和私营公司提供更多可用数据,促进大数据环境的良性发展。

4 我国的现状及启示

4.1 我国大数据伦理问题现状

大数据的强大张力使得各领域都更为重视社会发展对大数据的依赖,强调数据的社会和经济价值,但大数据环境下,数据非法获取和保存、数据滥用、数据主体对数据的控制权被削弱、数据垄断、数据偏向引导等事件的频繁发生也使得一些人开始关注到数据伦理问题。但目前来看,我国公民在大数据伦理方面也还没有形成明确认知。根据笔者的另一项调查发现,在 470 篇利用用户社交媒体数据进行研究的文献中,仅有一篇在文中有直接关于伦理的相关说明,研究者鲜有意识到其中可能存在伦理隐忧,但其中很多数据已经涉及到用户隐私。研究者对数据的采集,利用数据进行预测、关联分析,存储数据等均未获得用户的知情同意。我国的教育体制缺少这方面的系统训练,也导致我国公民相关认知和知识储备严重不足,这极有可能造成收益归个人,而风险留给社会,这将带来极大的科技风险[47]。这与国际上的研究发现存在共性,M. Zimmer 曾调查表示,学者利用 Twitter 数据进行研究的文献中,仅有 4% 提到了与研究设计和数据收集方法有关的伦理问题或者相关因素[48]。相较之下,我国这一问题则更加严峻,这也从一个侧面反应了提升我国公民数据素养的紧迫性。

从政策上看,在 2017 年 6 月 1 日,我国已正式施行《中华人民共和国网络安全法》,但这是一个原则性框架,无法全面覆盖当今复杂环境下的数据问题,还应该有很多配套制度和法规去完善。2019 年 1 月 23 日,中共中央网络安全和信息化委员办公室、工业和信息化工部、公安部、市场监管总局联合发布了《关于开展 APP 违法违规手机使用个人信息专项治理的公告》[49],规范对个人信息的收集和使用。2019 年 1 月 25 日,中共中央政治局举行第十二次集体学习时,习近平总书记进一步指出,要"规范数据资源利用,防范大数据等新技术带来的风险",李克强总理也多次强调,"对信息滥用、侵犯隐私……要依法依规进行打击清理"。2019 年 7 月 24 日,中央全面深化改革委员会第九次会议审议通过了《国家科技伦理委员会组建方案》,我国将加快建立科技伦理审查和风险评估制度[50]。可见,我国已经充分认识到科技发展中伦理建设的重要性,但相关规范建设也才刚刚起步,对大数据伦理问题的规制还有很长的路要走。

我国数据伦理规制的滞后,以及公民伦理意识的缺失无疑给数据利用带来了巨大的伦理风险,这种"缺位"也会制约我们未来的发展[51]。鉴于此,我国必须加快大数据利用中具体的伦理规范建设,注重对公民数据伦理意识的培养。

4.2 对我国的启示

4.2.1 明确大数据环境下个人权利

正如 EESC 调查所示,相关"意识"的缺乏是目前最紧迫的问题之一。当下,个人对数据的收集、传播和利用过程都呈现出一种"模糊"状态,因为个人权利的不明确导致对自身数据缺乏控制,同时自己作为数据利用者时也有可能侵犯他人隐私。2018 年 5 月实施的 GDPR 没有再将"个人权利"作为数据保护原则中的一条,而是用一章单独讨论,也足见个人权利的重要性。"使人们能够控制他们的数据"也成为欧洲一些国家数据伦理治理的关键目标之一[52]。我国在个人信息

chinaXiv:202304.00343v1

ChinaXiv合作期刊

保护方面尚未专门立法,对大数据环境下个人究竟享有怎样的数据权利更是没有明确规定。我国应加快制定《个人信息保护法》,对个人信息属性和保护模式给予重新定位,明确大数据环境下的个人权利体系,详细界定权利范围、权利内容等,使公众明白大数据环境下自己对自身的数据能够采取哪些操作,提高对个人数据的控制能力,提升数据提供者、数据使用者等各方的数据伦理意识。同时,也应逐步完善数据信息分类保护的法律规范,明确数据挖掘、存储、传输、发布以及二次利用等环节的权责关系[53]。也应注意强调个人信息保护不仅仅是为个人设定的权利,更是旨在构建一个平衡个人、信息利用者和社会利益的法律框架。

4.2.2 关注多利益主体的伦理需求

大数据环境下,每个人都在数据链条上同时扮演多个角色,在不同发展阶段,各主体数据需求不同,对大数据伦理问题的认知也不同。上述的一些调查结果表现出人们对大数据的态度存在内部冲突,EESC 的调查也体现了不同利益相关者对大数据伦理治理的不同态度。英国互联网安全和安保部长指出:人们对数据信任的看法存在很大差距,青少年就放弃了隐私并很乐意分享他们生活中的一切。但我们也有责任留意他们的利益,特别是在网上的权利方面[41]。不同的文化和群体拥有不同的隐私概念,针对不同的环境和对象所设定的隐私边界也不同。我国在面对大数据伦理问题时,也不能采取"一刀切"的形式,要防止"过分开发",但也要注意不能"过分保护"。保护个人可识别数据的需求也需要与这些数据可能被用于创造公共利益相平衡,根据我国的实际情况,将公益性价值取向考虑在内,尊重多元伦理需求,尽可能在数据开发利用中平衡效益和伦理。

chinaXiv:202304.00343v1

4.2.3 构建数字教育体系

人是技术使用的主体,也是引发伦理问题的关键[54],英国纳菲尔德生物伦理委员会主席 M. Jonathan 教授也表示:大多数违规都是人为的,而不是技术[55],人才建设是关键。2018 年,英国额外投入 4.06 亿英镑,用于数学、数字化和技术教育[56]。我国应借鉴欧盟经验,加大对数据教育的投入,针对不同的对象设计不同的课程体系。义务教育阶段,以科普性教育为主,使下一代掌握大数据基本常识,树立正确的"数据伦理"观念。大学阶段,设置系统的课程体系和培养计划,培养专业型数据人才。社会教育也应与学校教育联动,发挥图书馆的社会教育职能,举办专题讲座普及数据利用常识,指导、规范数据传播、利用行为,提高对数字社会的接纳程度,提升社会成员的数据素养。

4.2.4 建立数据使用者持续性审核责任机制

欧盟大数据伦理治理除了有益的经验值得借鉴外,在实践中面临的一些困境也值得我们参考。EESC 所倡导的建立欧洲健康电子数据库是通过规范数据利用来规避伦理问题,从源头降低数据的不规范利用,但却并没有获得认可。究其原因,主要是数据库存在数据泄露的风险。研究者一旦被授权可使用某些健康数据,即有可能将数据永久保存在其电脑中,这些数据日后可能有意或无意地被非法使用或超出其最初指定用途,由此引发的数据伦理问题有可能引发公众舆论而抵制研究。我国在鼓励数据开放利用的同时,应倡导建立对数据使用者的持续性审核责任机制。除了在数据申请阶段对数据使用者的资质、使用目的等进行详细的审核外,在进程中,也需定期就数据利用情况进行说明,特别是与申请时数据利用模式有变更的,需要提交申请并详细说明。并且,要明确权责关系,数据使用者必须对授权数据的合理使用和存储负责,一旦发生违法事件,需承担全部责任,并禁止再次申请利用数据。

5 结语

大数据时代,数据的采集、利用方式较以往发生了巨大变化,大数据技术是一种全新的信息技术,大数据伦理问题也是全新的问题,这给传统的法律规范、伦理道德带来了挑战,传统的隐私保护方式在大数据时代也基本失效。但数据价值的发挥不应该以伦理的丧失为代价,要有效平衡数据效益与伦理。我国应积极参考国际的治理实践,借鉴有益经验,从政策层面、行业层面、社会层面多管齐下,实现规范引导、严格监督。同时人们也应适时调整传统隐私观念和隐私领域认知,培育开放共享的大数据时代精神,使人们的价值理念与大数据技术发展以及文化环境更加契合,实现更有效的隐私保护。

参考文献:

[1] MARIAROSARIA T, LUCIANO F. How AI can be a force for good [J]. Science, 2018, 361(6404):751-752.
[2] FLORIDI L, TADDEO M. What is data ethics? [EB/OL]. [2019-09-14]. https://royalsocietypublishing.org/doi/pdf/10.1098/rsta.2016.0360.
[3] 梁芷铭. 大数据治理:国家治理能力现代化的应有之义[J]. 吉首大学学报(社会科学版), 2015, 36(2):34-41.
[4] 安宝洋. 大数据时代的网络信息伦理治理研究[J]. 科学学研究, 2015, 33(5):641-646.

[5] European Economic and Social Committee. The ethics of Big Data: Balancing economic benefits and ethical questions of Big Data in the EU Policy context[EB/OL]. [2019-08-14]. https://www.eesc.europa.eu/en/our-work/publications-other-work/publications/ethics-big-data.
[6] European Commission. Artificial Intelligence[EB/OL]. [2019-08-10]. https://ec.europa.eu/digital-single-market/en/artificial-intelligence.
[7] Genernal Data Protection Regulation[EB/OL]. [2019-11-14]. https://gdpr-info.eu/.
[8] European Commission. AI made in Europe[EB/OL]. [2019-08-09]. https://ec.europa.eu/digital-single-market/en/news/ai-made-europe.
[9] Ethics Guidelines for Trustworthy AI[EB/OL]. [2019-08-09]. https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai.
[10] DAVIS K, PATTERSON D. Ethics of big data[EB/OL]. [2019-11-14]. http://file.allitebooks.com/20150615/Ethics%20of%20Big%20Data.pdf.
[11] The Royal Social. Data governance: Public engagement review. 2017-6[EB/OL]. [2019-02-22]. https://royalsociety.org/~/media/policy/projects/data-governance/data-governance-public-engagement-review.pdf? la=en-GB.
[12] RICHARDS N, KING J. Big data ethics[J]. Social science electronic publishing, 2014, 1(2):1-6.
[13] Oxford Internet Institute. What is data ethics? [EB/OL]. [2019-03-10]. https://www.oii.ox.ac.uk/news/releases/what-is-data-ethics/.
[14] 薛孚, 陈红兵. 大数据隐私伦理问题探究[J]. 自然辩证法研究, 2015,31(2):44-48.
[15] 邱仁宗, 黄雯,翟晓梅. 大数据技术的伦理问题[J]. 科学与社会, 2014, 4(1):36-48.
[16] MOOR J. The future of computer ethics: you ain't seen nothin' yet! [J]. Ethics & information technology, 2001, 3(2):89-91.
[17] 郝庭帅. 当代社会生活的大数据化:困境与反思[J]. 社会发展研究, 2014,1(3):196-211.
[18] 陈仕伟. 大数据技术异化的伦理治理[J]. 自然辩证法研究, 2016,32(1):46-50.
[19] 上海社会科学院. 寻找"世界数据化"中的伦理平衡点[EB/OL]. [2019-03-20]. http://www.sass.org.cn/Default.aspx?tabid=167&ctl=Details&mid=607&ItemID=10767&SkinSrc=[L]Skins/shkxy/shkxy2_1.
[20] 王小杨. 大数据语境下的网络伦理治理转向[J]. 今传媒, 2015, 23(6):25-26.
[21] 陆伟华. 大数据时代的信息伦理研究[J]. 现代情报, 2014, 34(10):66-69.
[22] 赵素锦. 高科技时代的伦理风险及规避[J]. 求实, 2011,52(4):30-32.
[23] JUDD L. An approach to ethics in the information age[J]. Public relations review, 1995, 21(1):35-44.
[24] SEVERSOON R. The principles of information ethics[M]. New York: M. E. Sharpe, Inc., 1997.
[25] IEEE. IEEE ethics and member conduct[EB/OL]. [2019-02-25]. http://www.ieee.org/about/ethics.html.
[26] ACM. ACM code of ethics and professional conduct[EB/OL]. [2019-02-25]. http://www.acm.org/about-acm/acm-code-of-ethics-and-professional-conduct.
[27] Ethical guidelines for statistical practice[EB/OL]. [2019-02-25]. http://www.stats.uwo.ca/faculty/bellhouse/stat9544_2016_files/ASA%20EthicalGuidelines.pdf.
[28] Code of conduct for BCS members[EB/OL]. [2019-02-25]. http://www.bcs.org/upload/pdf/conduct.pdf.
[29] BERLEUR J. Ethics and the governance of the Internet: recommendations of IFIP - SIG9.2.2[M]//Computers and networks in the age of globalization. Boston: Springer, 2001.
[30] 舍恩伯格. 删除——大数据时代的取舍之道[M]. 袁杰,译. 杭州:浙江人民出版社,2013.
[31] ZWITTER A. Big data ethics[J]. Big data & society, 2014, 1(2):1-6.
[32] NUNAN D, DOMENICO M. Market research and the ethics of big data[J]. International journal of market research, 2013, 55(4): 505-520.
[33] 于英香. 大数据质疑述评[J]. 档案管理, 2015,32(1):17-19.
[34] 宋吉鑫. 网络伦理学研究[M]. 北京:科学出版社, 2012.
[35] 唐熙然. 大数据的伦理问题及其道德哲学——第一届全国赛博伦理学研讨会综述[J]. 伦理学研究, 2015,13(2):138-140.
[36] Unleashing the Potential of Cloud Computing in Europe[EB/OL]. [2019-03-24]. https://www.pdpjournals.com/docs/88053.pdf.
[37] An open, safe and secure Cyberspace[EB/OL]. [2019-03-24]. https://ec.europa.eu/home-affairs/what-is-new/news/news/2013/20130207_01_en.
[38] The Digital Skills and Jobs Coalition[EB/OL]. [2019-03-24]. https://ec.europa.eu/digital-single-market/en/digital-skills-jobs-coalition.
[39] Opinion 7/25 Meeting the challenges of big data[EB/OL]. [2019-03-24]. https://edps.europa.eu/sites/edp/files/publication/15-11-19_big_data_en.pdf.
[40] SWAN M. Philosophy of Big Data: Expanding the Human-Data Relation with Big Data Science Services[EB/OL]. [2019-08-14]. https://www.melanieswan.com/documents/Philosophy_of_Big_Data_SWAN.pdf.
[41] Big data-privacy principles under pressure[EB/OL]. [2019-03-25]. https://www.datatilsynet.no/globalassets/global/english/big-data-engelsk-web.pdf.
[42] FRED R. The Filter Bubble: What the Internet is hiding from you (review)[J]. Portal: libraries and the academy, 2011, 11(4): 1009-1011.

chinaXiv:202304.00343v1

[43] European Commission. Communication: building trust in human centric artificial intelligence [EB/OL]. [2019 - 08 - 11]. https://ec.europa.eu/digital-single-market/en/news/communication-building-trust-human-centric-artificial-intelligence.
[44] UK Public Research - Online Privacy [EB/OL]. [2019 - 03 - 25]. https://www.bigbrotherwatch.org.uk/wp-content/uploads/2015/03/Big-Brother-Watch-Polling-Results.pdf.
[45] The Direct Marketing Association [EB/OL]. [2019 - 03 - 26]. http://data.parliament.uk/writtenevidence/committeeevidence.svc/evidencedocument/science-and-technology-committee/big-data-dilemma/written/20453.pdf.
[46] European Commission. EU artificial intelligence ethics checklist ready for testing as new policy recommendations are published [EB/OL]. [2019 - 08 - 11]. https://ec.europa.eu/digital-single-market/en/news/eu-artificial-intelligence-ethics-checklist-ready-testing-new-policy-recommendations-are.
[47] 李侠. 科技伦理:没有约束的科技是危险的[N]. 光明日报, 2015 - 07 - 31(10).
[48] ZIMMER M, PROFERES N J. A topology of Twitter research: disciplines, methods, and ethics[J]. Aslib Journal of Information Management, 2014, 66(3):250 - 261.
[49] 中央网信办、工业和信息化部、公安部、市场监管总局关于开展App违法违规收集使用个人信息专项治理的公告[EB/OL]. [2019 - 08 - 14]. http://www.cac.gov.cn/2019-01/25/c_1124042599.htm.
[50] 孙敏. 守好科技伦理这道“门”[N]. 新华日报, 2019 - 08 - 13(1).
[51] 新华网[EB/OL]. [2019 - 02 - 21]. http://www.xinhuanet.com//fortune/2017-11/21/c_1121987357.htm.
[52] GOV.UK. Data Protection Act 2018 [EB/OL]. [2019 - 03 - 26]. https://www.gov.uk/government/collections/data-protection-act-2018.
[53] 杨维东. 有效应对大数据技术的伦理问题[N]. 人民日报, 2018 - 03 - 23(7).
[54] GOV.UK. AI Sector Deal [EB/OL]. [2019 - 03 - 26]. https://www.gov.uk/government/publications/artificial-intelligence-sector-deal/ai-sector-deal.
[55] The big data dilemma [EB/OL]. [2019 - 03 - 25]. https://publications.parliament.uk/pa/cm201516/cmselect/cmsctech/468/468.pdf.
[56] Industrial Strategy: building a Britain fit for the future [EB/OL]. [2019 - 03 - 26]. https://www.gov.uk/government/publications/industrial-strategy-building-a-britain-fit-for-the-future.

chinaXiv:202304.00343v1

The Practice of EU Big Data Ethics Governance and Its Enlightenment to China

Chen Yi
School of Information Management, Wuhan University, Wuhan 430072

Abstract: [**Purpose/significance**] Summarizing the experience of EU big data ethics governance and providing experience for big data ethics governance in China. [**Method/process**] This paper systematically sorted out the relevant policies of EU on the governance of big data ethics, analyzed the practical exploration of EU big data ethics governance, focused on the governance measures of the European Commission for Social Economics (EESC). On this basis, this paper summarized its governance experience, and summed up the enlightenment to the ethical governance of big data in China according to the actual situation of our country. [**Result/conclusion**] EESC starts from the human life cycle and summarizes the data ethics into 10 ethical issues. 5 governance measures are proposed from the perspectives of individuals, enterprises and research institutes. At the same time, it emphasizes to examine data ethics from multiple perspectives. The lag of the legal norms of data ethics and the lack of citizen's ethical awareness in China have brought great ethical risks to big data utilization. China should learn from EU, clarifying the individual rights in the big data environment, paying attention to the ethical needs of multi-stakeholders in the process of data governance, building a digital education system, and exploring to establish a researcher's continuous review responsibility mechanism.

Keywords: big data ethics governance